

03 OCT 2004

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2003 年 10 月 23 日 (23.10.2003)

PCT

(10) 国際公開番号  
WO 03/088209 A1(51) 国際特許分類:  
G10L 15/00,  
15/06, 15/28, G01C 21/00, G08G 1/0969KAISHA) [JP/JP]; 〒100-8310 東京都千代田区丸の内  
二丁目2番3号 Tokyo (JP).

(21) 国際出願番号: PCT/JP02/03680

(22) 国際出願日: 2002 年 4 月 12 日 (12.04.2002)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(71) 出願人 (米国を除く全ての指定国について): 三  
菱電機株式会社 (MITSUBISHI DENKI KABUSHIKI

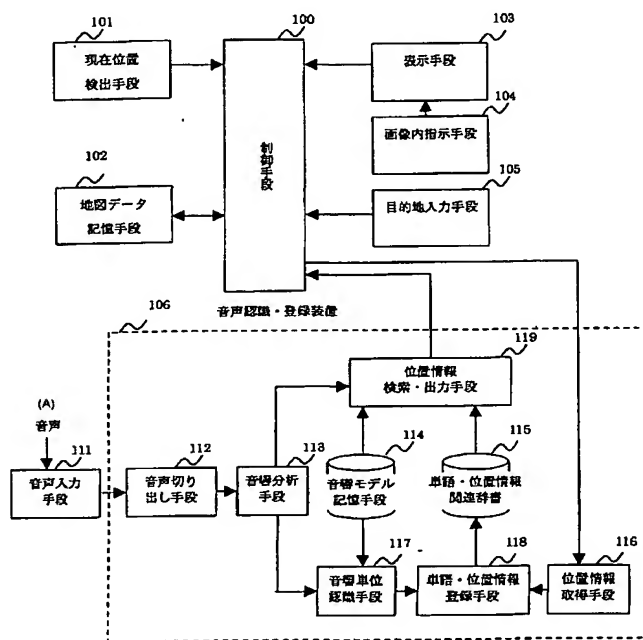
(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 丸田 裕三  
(MARUTA, Yuzo) [JP/JP]; 〒100-8310 東京都千代田区  
丸の内二丁目2番3号 三菱電機株式会社内 Tokyo  
(JP).(74) 代理人: 宮田 金雄, 外 (MIYATA, Kaneo et al.); 〒  
100-8310 東京都千代田区丸の内二丁目2番3号 三  
菱電機株式会社内 Tokyo (JP).

[続葉有]

(54) Title: CAR NAVIGATION SYSTEM AND SPEECH RECOGNIZING DEVICE THEREOF

(54) 発明の名称: カーナビゲーションシステム並びにその音声認識装置



101...PRESENT POSITION DETERMINING MEANS  
102...MAP DATA STORING MEANS  
100...CONTROL MEANS  
103...DISPLAY MEANS  
104...IN-IMAGE INSTRUCTING MEANS  
105...DESTINATION INPUTTING MEANS  
106...SPEECH RECOGNIZING AND REGISTERING DEVICE  
(A)...SPEECH  
111...SPEECH INPUTTING MEANS  
112...SPEECH SEGMENTING MEANS  
113...ACOUSTIC ANALYZING MEANS  
119...POSITION INFORMATION RETRIEVING AND OUTPUTTING MEANS  
114...ACOUSTIC MODEL STORING MEANS  
115...WORD AND POSITION INFORMATION RELATED DICTIONARY  
117...ACOUSTIC UNIT RECOGNIZING MEANS  
118...WORD AND POSITION INFORMATION REGISTERING MEANS  
116...POSITION INFORMATION ACQUIRING MEANS

(57) Abstract: A car navigation system having a speech recognizing function capable of retrieving position information through speech recognition without requiring any registration of speech recognition reference pattern of a word to be recognized. The car navigation system comprises an acoustic unit recognizing means (115) which collates an input acoustic feature vector produced by acoustic analysis of the speech uttered by a user with an acoustic model of each predetermined acoustic unit of a syllable or a phoneme and outputs a related acoustic unit recognition candidate series, a word and position information registering means (118) which relates the acoustic unit recognition candidate series to position information acquired from a car navigation system body to each other and registers them in a word and position information relation dictionary, and a position information retrieving and outputting means (119) which collates the input acoustic feature vector with the acoustic characteristic vector of each acoustic unit recognition candidate series in the word and position information relation dictionary, calculates the likelihood, and outputs to the car navigation system body the position information related to the acoustic unit recognition candidate series with the likelihood not less than a predetermined threshold.

BEST AVAILABLE COPY

[続葉有]



(81) 指定国 (国内): JP, US.

(84) 指定国 (広域): ヨーロッパ特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:

— 国際調査報告書

(57) 要約:

認識対象単語の音声認識標準パターンの登録を必要とせず音声認識により位置情報を検索できる音声認識機能付きカーナビゲーションシステムを提供する。

ユーザの発声する音声を音響分析した入力音響特徴ベクトルと、音節又は音素を音響単位とし所定の音響単位毎の音響モデルとを照合し、関連づけられた音響単位認識候補列を出力する音響単位認識手段(115)と、この音響単位認識候補列とカーナビゲーションシステム本体から取得した位置情報とを関連づけて単語・位置情報関連辞書に登録する単語・位置情報登録手段(118)と、上記入力音響特徴ベクトルと上記単語・位置情報関連辞書の各音響単位認識候補列の音響特徴ベクトルとを照合して尤度を計算し、この尤度が所定の閾値以上となる音響単位認識候補列に関連づけられた位置情報をカーナビゲーション本体に出力する位置情報検索・出力手段(119)とを備えた。

## 明 細 書

## カーナビゲーションシステム並びにその音声認識装置

## 技術分野

- 5 本発明は、音声認識装置を用いたカーナビゲーションシステムに係るものである。

## 背景技術

- カーナビゲーションシステムにおいては、走行中の操作  
10 の負荷を軽減し、特別な訓練を受けないユーザにとっても  
使いやすい操作インターフェースを実現するために、音声  
認識装置を搭載することが多くなっている。このよう  
なカーナビゲーション用音声認識装置では、一般に入力さ  
れた音声を特定するために、予め認識対象となる音声の周  
15 波数分布などを分析し、例えばスペクトルや基本周波数の  
時系列情報等の特徴として抽出し、そのパターンを対応さ  
せて格納する音声認識用辞書を備えている。そのための音  
声認識辞書としては、一般に不特定多数の話者を対象とし  
た音声の周波数分布などに対して統計的処理を行ったもの  
20 をパターン（以下、標準パターンとする）化し、そのパ  
ターンを単語と共に登録する不特定話者辞書及び、特定の話  
者を対象に入力音声のパターンを単語と共に登録する特定  
話者辞書がある。このような辞書は非常に多くの単語とそ  
の標準パターンを含んでいるため、外部記憶装置が記憶す  
25 る場合が多い。

このような音声認識装置においては、認識すべき音声

入力されると、外部記憶装置より上記両辞書に格納された各単語の標準パターンを順次読み込み、入力された音声の周波数パターンとこれらの標準パターンをパターンマッチングにより比較照合し、各単語に対する類似度を算出する。

- 5 次に算出された類似度が最も高い単語（パターンが最も近い単語）を、入力された音声であると認識し、その単語を出力するようにしている。つまり、入力された単語の周波数分布のパターンがどの単語パターンに最もよく似ているかを調べることによって、入力音声を判定している。

- 10 従来の音声認識装置および音声認識装置を利用したカーナビゲーションシステムは以上のように構成されているので、標準パターンを作成するために、音声認識される可能性のある単語の読み（例えば、「目的地設定」は「もくてきちせってい」と読むという情報）をあらかじめ記憶しておく必要があるという問題点があった。

- 15 また、別の従来の音声認識装置では、各単語の標準パターンのために費やすメモリを削減するため、登録時には全単語の標準パターンを一時的に作成して認識するが、そのうち認識が成功した単語についてのみ単語パターンを保持し、実際の使用時には保持している単語パターンのみについて照合を行う方式をとるものもある。

- 20 また、従来の音声認識装置および音声認識装置を利用したカーナビゲーションシステムは以上のように構成されているので、認識が成功した単語でなければ、特定辞書に単語パターンを標準パターンとして登録することができない。
- 25 このため、認識が成功することのない単語の読み方、すな

わち、あらかじめ用意された単語の読み以外の読み方では登録できないという問題点があった。

また、従来の音声認識装置および音声認識装置を利用したカーナビゲーションシステムは以上のように構成されているので、単語を認識するためには少なくとも一度はその単語を使用している必要があり、それまで一度も使用されていない単語を認識することができないという問題点があった。

また、従来の音声認識装置及び音声認識装置を利用したカーナビゲーションシステムは以上のように構成されているので、外部記憶装置が記憶する大量の単語とその標準パターンを順次読み出して、照合処理を行うために処理時間を要し、応答性能を得にくいという問題点があった。

本発明は、上述のような課題を解決するためになされたもので、ユーザの発声した音声標準パターンとして登録するようにすることで、予め音声認識をする必要のある単語の読みを記憶する必要性を廃し、標準的な読み以外の読みをも認識可能とすることを目的とする。

また、音響単位混同行列を用いて登録されたユーザの音声とそれに音響上類似する語を展開し、この類似語をあわせて登録することで、検索時に発声されたユーザの音声と登録されたユーザの音声の認識照合精度を高めることを目的とする。

さらに、外部記憶装置中に記憶された単語とその読みをも用いて照合する場合には、読みの類似する単語をグルーピングし、各グループの代表単語とユーザの発声した音声

との認識照合を行い、その尤度が所定の閾値以上となる場合のみユーザの発声した音声とそのグループに属する他の単語との照合を実行することで、無駄な照合処理を廃し、また外部記憶装置の単語と読みをグループ単位でメモリに  
5   ロードすることでメモリの使用効率を向上することを目的とする。

#### 発明の開示

第1の発明に係る音声認識装置は、音声入力手段により  
10   入力されたユーザの発声する音声を音響分析して入力音響特徴ベクトルを出力する音響分析手段と、音節又は音素を音響単位とし所定の音響単位毎の音響モデルを予め記憶する音響モデル記憶手段と、この音響モデルと上記入力音響特徴ベクトルとを照合し音響単位認識候補列を出力する音  
15   響単位認識手段と、この音響単位認識候補列とカーナビゲーションシステム本体から取得した位置情報とを関連づけて単語・位置情報関連辞書に登録する単語・位置情報登録手段と、上記音響分析手段により出力された入力音響特徴ベクトルと上記単語・位置情報関連辞書の各音響単位認識  
20   候補列の音響特徴ベクトルとを照合して尤度を計算しこの尤度が所定の閾値以上となる音響単位認識候補列に関連づけて単語・位置情報関連辞書が記憶する位置情報をカーナビゲーション本体に出力する位置情報検索・出力手段とを備えるものである。

25   第2の発明に係る音声認識装置は、上記音響分析手段の認識精度により人間が発声する発声音響単位が異なる認識

- 結果として認識される確率を各認識結果毎に予め記憶する音響単位混同行列記憶手段と、上記音響単位認識手段が出力する音響単位認識候補列の各音響単位ごとに上記認識確率が一定閾値以上となる認識結果音響単位からなる認識結果音響単位リストを抽出する音響単位混同行列検索手段と、
- 5 上記音響単位認識候補列の各音響単位をその認識結果音響単位リストの認識結果音響単位に置換することにより得られる上記音響単位認識候補列の類似候補を出力する単語展開手段とを備え、
- 10 上記単語・位置情報登録手段が、この類似候補をカーナビゲーションシステム本体から取得した位置情報とを関連づけて上記単語・位置情報関連辞書に登録する構成としたものである。

- 第3の発明に係る音声認識インデックス検索装置は、複数の単語をその読みが類似するものにグループ化しそのグループの中の一つの単語を代表単語としてこの代表単語とそのグループの関連を記憶する類似単語インデックス手段と、上記類似単語インデックス手段の記憶するグループ毎の代表単語の音響特徴ベクトルと所定の音響特徴ベクトル
- 15 とを照合してそれぞれの尤度を計算しそのうち所定の閾値以上となる尤度の代表単語が属するグループの各単語の音響特徴ベクトルと上記所定の音響特徴ベクトルとをそれぞれ照合して尤度を計算し最も尤度が高い単語を出力するグループ内類似単語検索手段とを備えるものである。
- 20

- 25 第4の発明に係る音声認識インデックス作成手段は、複数の単語からなる元集合から一つの単語を代表単語として

選択する代表単語選択手段と、この代表単語の音響特徴ベクトルとの尤度が所定の閾値となる音響特徴ベクトルを有する単語を上記元集合から抽出して代表単語と同じグループに含める音響類似単語グルーピング手段と、このグループに属する単語を元集合から除いた単語の集合を上記代表単語選択ステップの処理する元集合としてこの代表単語選択ステップに引き渡す元集合置換手段とを備えるものである。

第5の発明に係る音声認識装置は、上記位置情報検索・出力手段が、第3の発明又は第5の発明に係る音声認識インデックス検索装置を備え、この位置情報検索・出力手段を、上記単語・位置情報関連辞書と外部記憶装置とが記憶する単語とその読みと位置情報とをこの音声認識インデックス検索装置により検索し出力する構成としたものである。

第6の発明に係る音声認識装置は、上記単語・位置情報登録手段が、上記音響分析手段の認識精度により上記音響単位認識候補列の各音響単位がこの音響単位に対応する上記音響単位認識候補列の類似候補の音響単位として認識される確率を上記音響単位認識候補列の音響単位毎に上記音響単位混同行列記憶手段より読み込み音響単位毎の確率を含む上記類似候補の確率リストを作成して上記類似候補とこの確率リストとを上記単語・位置情報関連辞書に登録するとともに、

上記位置情報検索・出力手段が、上記単語・位置情報関連辞書が記憶する類似単語候補とこの類似単語の確率リストを読み込みこの確率リストの確率が所定の閾値以上の場合



は上記入力音響特徴ベクトルと上記音響特徴ベクトル生成手段が出力した音響特徴ベクトルとを照合して尤度を計算しこの尤度が所定の閾値以上となる音響単位認識候補列を出力し、

- 5 この確率が上記所定の閾値未満の場合は外部記憶装置が記憶する単語とその読みと位置情報を第3の発明又は第5の発明に係る音声認識インデックス検索装置により検索する構成としたものである。

- 10 第7の発明に係るカーナビゲーションシステムは、現在位置検出手段と、地図データ記憶手段と、画像表示手段と、画像内指示手段と、目的地入力手段と、音声認識装置とを備え、この音声認識装置を第1の発明又は第2の発明又は第6の発明又は第7の発明のいずれか一の音声認識装置とするものである。

15

図面の簡単な説明

図1は本発明の実施例1を示すカーナビゲーションシステム用音声認識登録装置の構成図、

- 20 図2は本発明の実施例1を示すカーナビゲーションシステム用音声認識登録装置に係る処理のフローチャート、

図3は本発明の実施例2を示すカーナビゲーションシステム用音声認識登録装置の構成図、

図4は本発明の実施例2を示すカーナビゲーションシステム用音声認識登録装置に係る処理のフローチャート、

- 25 図5は本発明の実施例2における音響単位混同行列の例を示す図、

図 6 は本発明の実施例 2 におけるカーナビゲーションシステム用音声認識登録装置において、類似している音節列を作成する処理を示すフローチャート、

5 図 7 は本発明の実施例 2 におけるカーナビゲーションシステム用音声認識登録装置において、音響単位認識結果の音響単位を類似候補によって置換することにより得られる展開単語の例を示す図、

図 8 は本発明の実施例 3 におけるカーナビゲーションシステム用音声認識登録装置の構成図、

10 図 9 は本発明の実施例 3 におけるカーナビゲーションシステム用音声認識登録装置に係る処理を示すフローチャート、

図 10 は本発明の実施例 4 における音声認識インデックス検索装置の構成図、

15 図 11 は本発明の実施例 4 における音声認識登録装置による処理のフローチャート、

図 12 は本発明の実施例 5 における音声認識インデックス作成方法を示す構成図、

20 図 13 は本発明の実施例 5 における音声認識インデックス作成方法による処理を示すフローチャートである。

発明を実施するための最良の形態

以下に、本発明に係るカーナビゲーションシステム用音声認識装置を各実施例に従って図面に基づいて説明する。

25

実施例 1 .

図 1 は、本発明の実施例 1 の音声認識装置付きナビゲーションシステム装置のシステム構成を示したものである。図において、101 は車両の現在位置を検出する現在位置検出手段であって、周知構成の地磁気センサ、ジャイロスコ  
5 コープ、距離センサ及び衛星からの電波に基づいて車両の位置を検出する GPS (Global Positioning System) のための GPS 受信機を有している。制御手段は、これらの各センサが性質の異なる誤差を有しているため、各々補間しながら使用する。

10 102 は地図データ記憶手段であって、位置検出の精度向上のための所謂マップマッチング用データ、地図データ及び目印データを含む各種データを記憶している。地図データ記憶手段 102 が用いる記憶媒体としては、CD-ROM を用いるのが一般的であるが、DVD-ROM、メモ  
15 リカードなどの他の媒体を用いるようにしてもよい。

103 は画面を表示するための表示手段であって、一般的にはカラー表示機能を備えており、その画面には、現在位置検出手段 101 から入力された車両現在位置マークと、地図データ記憶手段 102 より入力された地図データと、  
20 さらに地図上に表示する誘導経路や後述する設定地点の目印などの付加データとを重ねて表示するようになっている。

104 は画像内指示手段であって、ポインタなど、表示手段によって表示された画像の一部を特定する指示手段である。

25 105 は目的地の位置や名称を入力する目的地入力手段であって、例えばリモートコントロール端末（以下、リモ

ゴンと称する) に対する操作を受信するリモコンセンサや操作スイッチ群により実現され、目的地の位置や名称を入力するために用いるが、その他カーナビゲーションシステムに対して指示を入力するためにも用いられる。操作スイッチ群は、例えば、表示手段と一体になったタッチスイッチ若しくはメカニカルスイッチなどが用いられる。

1 1 1 はマイクなど音声を取り込む音声入力手段であって、音声をデジタル信号に変換する。1 1 2 は音声切り出し手段であって、音声入力手段 1 1 1 が取り込んだ音声から、無音区間やノイズ区間を除き、人間の発声している範囲のみ検出する手段である。1 1 3 は音響分析手段であって、音声切り出し手段 1 1 2 が切り出した音声から、音声認識に適するような特徴を抽出する手段である。

1 1 4 は音響モデル記憶手段であって、各音響単位 (たとえば、日本語では、音節 (あ/a/、い/i/、う/u/、え/e/、お/o/、か/ka/、き/ki/、く/ku/、け/ke/、こ/ko/、など)) とその音響パターンを記憶している。1 1 5 は単語・位置情報関連辞書であって、認識対象となる単語とその音響パターンとこの単語に関連する位置情報が登録されるが、初期状態 (工場出荷時) にはなにも登録されていない。1 1 6 は位置情報取得手段であって、カーナビゲーションシステム本体より位置情報を取得し出力する手段である。1 1 7 は音響単位認識手段であって、音響分析手段 1 1 3 が抽出した入力音声の特徴と、音響モデル記憶手段が記憶している音響単位の音響パターンを照合して連続音声認識を行い、尤度 (確からしさ) の最も高い音響単位の列を認識候補と

- して出力する手段である。118は単語・位置情報登録手段であって、音響単位認識手段117が出力した認識候補を位置情報取得手段116から取得した位置情報とともに単語・位置情報関連辞書115に登録する手段である。1
- 5 19は位置情報・検索出力手段であって、単語・位置情報関連辞書115に登録されている各単語と音響モデル記憶手段が記憶している音響単位の音響パターンから各単語の音響特徴ベクトルを生成し、音響分析手段113が抽出した入力音声の音響特徴ベクトルとの尤度を計算した上で、
- 10 単語・位置情報関連辞書115の記憶する単語の中から入力音声とこの尤度の高い単語の上位N個（Nは整数）を認識候補とし、単語・位置情報関連辞書115がこの認識候補に関連づけて記憶している位置情報をカーナビゲーションシステム本体に出力する手段である。
- 15 次に動作を図2を用いて説明する。ここでは、日本語を例にとり、音響単位を音節として説明する。最初に、単語を登録する場合の動作について説明する。まずユーザは、目的地入力手段105に付属する操作スイッチを用いて、登録処理の指示を入力する（ステップS100：YES）。
- 20 次にユーザが登録したい位置（たとえば自宅など）を目的地入力手段105により入力すると（ステップS201）、表示手段103は入力した目的地を含む地図画面を表示する（ステップS202）。次にユーザは、画像内指示手段104によって登録したい位置（この場合、自宅）を画面上
- 25 に指示する（ステップS203）。次にユーザはその名称を発声（たとえば「自分の家」など）する（ステップS20

4)。

音声入力手段 1 1 1 は入力された音声に対してサンプリングを行い、デジタルデータに変換して、このデータを音声切り出し手段 1 1 2 に送付する。音声切り出し手段 1 1 2 は、入力されたデータのうち、入力された音声のパワー情報などを利用して無音区間やノイズ区間を除き、実際に音声である部分を切り出す。次に音響分析手段 1 1 3 は音声切り出し手段 1 1 2 が切り出した区間の音声データから、音響認識に適するような特徴（たとえば周知であるメルケプストラムなどの特徴）を抽出する（ステップ S 2 0 5）。次に音響単位認識手段 1 1 7 は、音響モデル記憶手段 1 1 4 が記憶している各音響単位（この場合は音節）の特徴と、音響分析手段 1 1 3 が出力した特徴から連続音声認識を行い、尤もらしい音響単位列（この場合は音節列）を認識候補として出力する。この場合には認識性能が良好であれば、「じぶんのいえ」という音節列が認識候補になる（ステップ S 2 0 6）。

次に位置情報取得手段 1 1 6 は、画像内指示手段 1 0 4 によって指示されている地図上の位置を地図データ記憶手段 1 0 2 から取得し、これを出力する。例えばこの場合では、自宅の住所の位置情報（東経・北緯や番地など）を出力する。次に単語・位置情報登録手段 1 1 8 は、音響単位認識手段 1 1 7 が出力した認識候補の音節列を、位置情報取得手段 1 1 6 が出力した位置情報とともに単語・位置情報関連辞書 1 1 5 に登録する。この場合には、「じぶんのいえ」という音節列が自宅の住所（例えば「鎌倉市大船 5 -

1 - 1」)とともに単語・位置情報関連辞書 1 1 5 に登録される。

以上の処理を繰り返すことによって、単語を登録する。

次に、カーナビゲーションシステムによる住所検索の動作について説明する。まずユーザは検索処理の指示を入力する(ステップ S 1 0 0 : N O)。次にユーザは検索・表示させたい住所(たとえば「自分の家」)を発声する(ステップ S 1 0 1)。

音声入力手段 1 1 1 は、入力された音声に対してサンプリングを行い、デジタルデータに変換して、音声切り出し手段 1 1 2 に送付する。音声切り出し手段 1 1 2 は、入力されたデータのうち、入力された音声のパワー情報などを基準にして無音区間やノイズ区間を除き、実際に音声である部分を切り出す。次に音響分析手段 1 1 3 は、音声切り出し手段 1 1 2 が切り出した区間の音声データから、音声認識に適するような特徴(たとえば周知であるメルケプストラムなどの特徴)を、音響特徴ベクトルとして抽出する(ステップ S 1 0 2)。次に、位置情報検索・出力手段 1 1 9 は、単語・位置情報関連辞書 1 1 5 に格納された単語の音響単位に基づいて、音響モデル記憶手段 1 1 4 が記憶する音響パターンを合成し、単語・位置情報関連辞書 1 1 5 の音響特徴ベクトルを生成し、この音響特徴ベクトルと音響分析手段 1 1 3 が出力した音響特徴ベクトルとを照合して単語認識を行い、尤度の高い単語を認識結果単語とする。この場合には、「じぶんのいえ」という単語が認識結果単語となる(S 1 0 3)。

次に位置情報検索・出力手段 1 1 9 は、認識結果単語の尤度が所定閾値より以上かどうかを検定し、所定閾値未満の場合（ステップ S 1 0 4 : N O）には認識が失敗したとして、エラー表示をし、音声入力待ち状態になる。認識結果単語の尤度が所定閾値より以上の場合には（ステップ S 1 0 4 : Y E S）、単語・位置情報関連辞書 1 1 5 が認識結果単語と関連づけて格納する位置情報を取得し、カーナビゲーション本体に出力する。この場合には「鎌倉市大船 5 - 1 - 1」という位置情報が出力される。次に表示手段 1 0 3 は、取得された位置情報が示す位置が中央になるように、地図情報を画像表示させる。

以上の処理によりユーザは、自分の自由な読み方による目的地の呼び名の音声を登録し、また音声により検索することが可能になる。

また、本実施例は以上のように構成されているので、工場出荷時において予め住所の読みを準備する必要がなく、またこのことにより、ユーザによって使用される見込みがないにも関わらず工場出荷時には一通り揃えておく必要があった音声データを廃することができるので、その音声データのために用いていた記憶容量を削減できるという効果がある。

また、上記位置情報検索・出力手段 1 1 9 は、単語・位置情報関連辞書に登録された単語に対してのみ認識処理を行うので、検索対象となる単語数が一般的に少なく、類似単語も少ないことから、認識が高速かつ高精度であるという効果がある。



また、本実施例では、登録処理において、尤度の高い音節列を1つのみ登録するようにしたが、尤度の高い音響単位列を上位N個(Nは整数)を登録することにしてもよい。この場合、例えば「自分のいえ」という発声に対して、第  
5 1位の音節列「じぶんのいえ」、第2位の音節列「じぶんのゆえ」、などを同じ位置情報に関連づけて単語・位置情報関連辞書に登録する。また所定の閾値を超える尤度を有する音響単位列を単語・位置情報関連辞書に登録するようにしても同様である。このように構成すれば、位置情報検索を行  
10 う際の単語認識において、発声のあいまい性を許容するカーナビゲーションシステムを提供することができる。

また、ここでは日本語による音声認識について説明したため、日本語の最小発声単位である音節について説明したが、外国語の場合は、子音などを考慮して、例えば音素な  
15 どの最小の発声単位での認識手段と、音響モデルを用いれば、同様な実施例が構成できることはいうまでもない。

## 実施例2.

次に本発明の別の実施例に係る音声認識装置について説明する。実施例2は、登録時の認識結果に誤りが含まれて  
20 いても、認識性能が劣化しにくい音声認識装置を提供するものである。図3は本実施例に係る音声認識装置をナビゲーションシステム装置に適用した場合のシステム構成を示したものである。図において、201は単語展開手段であり、認識結果から類似候補を展開するものである。また、  
25 202は音響単位混同行列記憶手段であり、音響単位混同

行列を記憶している。また、図 4 と図 6 は処理の流れを説明する図、図 5 と図 7 は説明のための図である。ここでは、日本語を例にとり、音響単位を音節として説明する。

- 音響単位混同行列記憶手段 202 が記憶する音響単位混  
5 同行列（この場合は音節混同行列）は、図 5 に示すように、認識処理上混同しやすい音響単位（この場合は音節）の対が行列の形で表現されている。各行列の要素は、条件付き確率で表現されている。実際の発声が X であり、この X の認識結果が Y となる条件付き確率を  $P(X|Y)$  と表すこととする。たとえば認識結果が音節「か (/ka/)」であっても、  
10 実際の発声は「た (/ta/)」であった確率  $P(ta|ka)$  と表す。この条件付き確率は、事前に大量データによって計算される。事前に大量データを用いた認識により、音節「か (/ka/)」である認識結果が N 個あり、このうち音節「た (/ta/)」を  
15 実際の発声とするものが M 個あった場合には、 $P(ta|ka) = M/N$  として計算する。理想的には、実際の発生が「か (/ka/)」である場合、認識結果の音節も「か (/ka/)」であるべきなので、 $P(ka|ka)$  は 1 となり、その他の成分（例えば  $P(ta|ka)$ ）は 0 となる。しかし実際には、認識性能の  
20 限界などにより、そうとはならない。これらの値は多人数かつ大量のデータによって事前に計算をしておく。以降、 $P(ka|ka)$  など、認識結果音節と実際の発声音節が一致している条件付き確率を、音響単位混同行列の対角成分と呼び、 $P(ta|ka)$  など、認識結果音節と実際の発声音節が一致していない条件付き確率を、音響単位混同行列の非対角成分と呼ぶ。  
25

次に動作を説明する。ステップ S 2 0 1 乃至ステップ S 2 0 6 は、実施例 1 と同じであるので説明を省略する。次に単語展開手段 2 0 1 は、音響単位認識手段 1 1 7 が出力した音節列をもとに、それと認識処理上類似している音響  
5 単位列（この場合は音節列）を作成する（ステップ S 3 0 1）。次に単語・位置情報登録手段 1 2 0 は、音響単位認識手段 1 1 7 が出力した音響単位列（音節列）と、単語展開手段が出力した音響単位列を、位置情報取得手段 1 1 6 が取得した位置情報とともに単語・位置情報関連辞書 1 1 5  
10 に登録する（ステップ S 3 0 2）。

類似している音節列は、図 6 に示す処理によって作成される。まず変数  $i$  を 1 に初期化する（ステップ S 4 0 1）。次に音響単位認識手段 1 1 7 が出力した音節列の第  $i$  番目の音節を取得する。この取得した音節を  $a$  とする（ステップ S 4 0 2）。次に或る音節  $b$  ( $\neq a$ ) について、音響単位  
15 混同行列記憶手段 2 0 2 が記憶している音響単位混同行列を参照して、条件付き確率  $P(b|a)$  を取得する（ステップ S 4 0 3）。その後、 $P(b|a)$  が所定閾値より大きいかどうかを調べ、大きい場合には第  $i$  音節を置き換える代替候補として記憶しておく（ステップ S 4 0 4、ステップ S 4 0 5）。  
20 以上の処理をすべての音節  $b$  ( $\neq a$ ) について行ない（ステップ S 4 0 6）、さらに変数  $i$  をインクリメントして音節列の各音節について行う（ステップ S 4 0 7、ステップ S 4 0 8）。次に第  $i$  音節を認識結果のままの場合と代替候補  
25 にした場合の音節列をすべての組み合わせで作成し、それらを類似単語とする。

図 7 に処理例を示す。例えば発声が「goruhujoo」（ゴルフ場）であるにもかかわらず、音節認識手段の性能の限界により、認識音節列が「gorucuzoo」であったとする。このとき、 $P(hu|cu)$ 、 $P(u|cu)$ 、 $P(jo|zo)$  がそれぞれ所定閾値より大きいので、第 3 音節の代替候補として  $hu$  と  $u$  が、第 4 音節の代替候補として  $jo$  が記憶される。認識結果の音節と代替候補のすべての組み合わせにより類似候補を作成すると、類似候補は、「gorucuzoo」、「goruhuzoo」、「goruuzoo」、「gorucujoo」、「goruhujoo」、「goruujoo」の 6 つとなる。これらの類似候補について、同じ位置情報に関連づけて単語・位置情報関連辞書 1 1 5 に登録する。

住所検索は、実施例 1 と同じ処理を行うが、検索時に「ゴルフ場」と発声した場合、上記の 6 つの単語のいずれかが認識結果となれば、ゴルフ場の位置が正しく表示される。

実施例 2 は以上のように構成されているので、単語登録時の認識結果に誤りが含まれていても、認識性能が劣化しにくい音声認識装置を提供することができる。

また、実施例 2 では、類似単語の作成において、尤度の高い音節列についてのみ類似単語を作成するようにしたが、尤度の高い音節列を上位  $N$  個（ $N$  は整数）について、それぞれの類似単語を作成するようにしてもよい。このように構成すれば、単語認識の際に、発声のあいまい性などをさらに許容する音声認識装置を提供することができる。

また、ここでは日本語による音声認識について説明したため、日本語の最小発声単位である音節について説明したが、外国語の場合は、子音などを考慮して、例えば音素な

どの最小の発声単位での認識手段と、認識辞書、混同行列を用いれば、同様な実施例が構成できることはいうまでもない。

### 5 実施例 3 .

次に本発明のさらに別の実施例に係る音声認識装置について説明する。実施例 3 は、検索時に単語辞書に登録されていない単語を発声した場合であっても、この単語に関連づけられた情報を検索できる音声認識装置を提供するものである。図 8 は本実施例に係る音声認識装置付きナビゲーションシステムのシステム構成を示したものである。図において、121 は外部記憶手段であり、例えば DVD などの大容量記憶手段によって構成され、住所など検索対象となる単語と読みと対応する位置情報が記憶されている。122 は音声認識インデックス装置であり、これは外部記憶手段に格納されている単語と読みを部分的に読み込んで記憶するための補助的記憶装置であって、通常は、外部記憶手段に格納されている単語と読みを一度に格納できる容量の数分の 1 以下のメモリが割り当てられている。

次に動作を図 9 を用いて説明する。登録時の処理（ステップ S 2 0 1 ～ステップ S 2 0 8）は実施例 1 と同じであるので説明を省略する。検索時の処理でステップ S 1 0 1 乃至ステップ S 1 0 4 は実施例 1 と同じであるため、説明を省略する。次に単語認識手段 1 1 4 は、各認識候補の尤度がすべて所定閾値より小さい場合（ステップ S 1 0 4 : NO）、発声された単語がまだ単語・位置情報関連辞書 1 1

5 に登録されていないと判断する。この場合は、音声認識インデックス装置 1 2 2 に対して入力音声の音響特徴ベクトルを与えて、外部記憶手段 1 2 1 にある住所の単語と読みの情報と位置情報の検索を要求する(ステップ S 5 0 1)。

- 5 音声認識インデックス装置は、外部記憶装置の内容から入力音声の音響特徴ベクトルと単語の読みが一致する単語を検索して、その単語に関連づけられる位置情報を検索結果とするので、その検索結果が成功かどうかを検定する(ステップ S 5 0 2)。この結果が成功である場合は、表示手段
- 10 1 0 3 は取得した位置情報の位置を画面の中央に表示させる。検索結果が失敗(該当する情報なし)の場合(ステップ S 5 0 2 : N O)、制御手段は認識エラーとして認識できなかった旨の表示を表示手段により行う。

- 15 以上の処理により、まだ登録処理を行っていない単語についても、外部記憶装置を参照するので、認識処理が可能となる。

#### 実施例 4 .

- 次に、上記音声認識インデックス装置の実施例について
- 20 説明する。この音声認識インデックス装置は、音声認識対象となる単語が大量にある場合に、メモリ使用量を抑制しながら高速に音声認識を行うものである。図 1 0 は実施例 4 に係るシステム構成を示したものである。図において、
- 3 0 1 はグループ内類似単語検索手段である。3 0 2 は類似単語インデックス手段、3 0 3 は代表単語テーブルであり、この音響認識インデックス装置が検索対象とする複数
- 25

の単語の中から選択した単語を保持するテーブルである。

304は単語グループ群であり、検索対象となる複数の単語の中から代表単語テーブル303の各代表単語に音響的に類似する単語同士を分類したグループの集合である。

- 5 次に動作を図11を用いて説明する。図11では、類似単語インデックス手段302が記憶する代表単語配列の*i*番目の要素を表現するために、代表単語(*i*)という表記を用いている。まず*i* = 1として*i*を初期化する(ステップS601)。次に代表単語(*i*)の音響特徴ベクトルを生成する(ステップS602)。この音響特徴ベクトルは電源投入時や検索処理時毎に生成する必要は必ずしもなく、代表単語の読み情報として予め与えられたものを用いてもよい。また電源投入時や検索処理時毎に生成する場合は音素や音節に基づく音響モデルと代表単語の音素や音節とから
- 10 音響特徴ベクトルを合成してもよい。
- 15

- 次に、代表単語(*i*)の音響特徴ベクトルと検索キーとなる入力音響特徴ベクトルとの尤度を計算し(ステップS603)、この尤度が所定の閾値以上か否かを検定する(ステップS604)。この尤度が所定の閾値以上である場合
- 20 (ステップS604: YES)は、類似単語インデックス手段が代表単語(*i*)に関連づけて記憶するグループの中から、入力音響特徴ベクトルとの尤度が最大の単語となる単語を求める(ステップS605)。ステップS604において尤度が所定の閾値未満の場合(ステップS604: N
- 25 O)は*i*に1を加え(ステップS607)、次の代表単語の評価を行う。

なおステップ S 6 0 5 を実行する時点において、代表単語（i）に関連するグループが外部記憶装置上にあり、まだメモリ上に存在していない場合は、このグループに属する単語と読みとその単語に関連する情報（位置情報など）をメモリに読み込む処理を実行してもよい。またその一方で、すでに入力音響特徴ベクトルとの尤度が所定の閾値以上にはならなかった代表単語が存在する場合は、その代表単語に関連するグループ内の単語を記憶していたメモリを解放することによって、この検索処理に使用するメモリの総量を抑制することも可能となる。

また、代表単語テーブル 3 0 3 の保持する各単語は、事前に所定の単語としておくこともできるし、また後述するように検索対象となる複数の単語の中から自動的に決定することも可能である。さらに、代表単語テーブル 3 0 3 はこの音響認識インデックス装置の電源投入時に作成してもよいし、外部記憶装置に予め記憶させておいた所定の代表単語テーブルを読み込むように処理してもよい。

また、単語グループ群 3 0 4 が含む個々のグループは、論理的な概念であるため、物理的にまとまって存在している必要はなく、例えば代表単語テーブル 3 0 3 にそのグループの代表単語とともにそのグループに属する単語と読みと関連づけられた情報の存在場所に対するポインタやアドレスのみを保持する形態で処理することも可能である。

また実施例 4 では、外部記憶装置は、カーナビゲーション装置の一部であるものとして説明したが、無線によるデータ転送手段を介して、他の装置に存在する記憶装置をア



クセスしても同等な効果を奏する。この場合、代表単語テーブルも無線通信によって取得する方法も考えられる。

#### 実施例 5 .

- 5      次に、複数の単語を音響的に類似する単語のグループに分類し、その中から代表単語を選択する音声認識インデックス作成手段の実施例について説明する。上記音響認識インデックス装置は、複数かつ任意の単語を含む単語の集合の中から入力音響ベクトルに基づいて一致する単語を効率的に検索する手段を提供するが、この音響認識インデックス装置によって効率的に単語の検索を実行するためには、
- 10      代表単語同士が相互に音響的に非類似であって、かつ検索対象となる複数の単語が必ずいずれかの代表単語に音響的に類似し、その代表単語のグループに属する必要がある。
- 15      この音声認識インデックス作成装置は、複数単語を相互に十分尤度が離れた複数のグループに分類することを目的とするものである。

- 図 1 2 は本発明の音声認識インデックス作成方法の実施例に係る構成を示したものである。図 1 2 において、
- 20      4 0 1 は、音声認識の対象となる複数の単語を含む元集合から、代表単語を選択する代表単語選択手段である。代表単語選択手段 4 0 1 が選択する代表単語は、元集合が含む複数の単語のうちの任意の一つの単語でよい。4 0 2 は音響類似単語グルーピングステップであって、元集合が含む複数の
- 25      単語から代表単語選択手段 4 0 1 が選択した代表単語以外の単語のうち、読みがこの代表単語と類似する単語を選択

する。ここでは、代表単語の音響特徴ベクトルと元集合に含まれる各単語の音響特徴ベクトルとの尤度を計算し、この尤度が所定の閾値以上となる単語を代表単語に関連するグループに属するものと判定する。403は元集合置換手段であって、音響類似単語手段によって抽出された上記代表単語に読みが類似する単語を元集合から取り除き、この状態の元集合をもとの元集合に置換し、上記代表単語選択手段401が再度処理可能とする。

次に、動作を図13を用いて説明する。最初に元集合の単語の個数を検定する(ステップS801)。元集合に含まれる単語の個数が0である場合は、もはや代表単語の選定も単語のグルーピングも必要がないことになるため、処理を終了する(ステップS801:NO)。一方、元集合の単語の個数が0でない場合は、代表単語の設定と単語のグルーピングを行う必要があるから、次のステップに進む(ステップS801:YES)。

次に、元集合の単語のリストの先頭要素を代表単語とする(ステップS802)。なお本実施例では、処理を最も簡単にするために元集合の単語のリストの先頭要素を機械的に代表単語としているが、代表単語とは音響的に相互に類似する単語の集合のどの要素でもよく、またこの後に続くステップで代表単語に音響的に類似する単語を事後的に元集合の中から収集していくことになるので、この時点における代表単語は元集合のどの要素でもよい。

次に、代表単語の音響特徴ベクトルを求める(ステップS803)。元集合に含まれる単語が外部記憶装置上にある

場合は、各単語の読みに替えて、あるいは各単語の読みとともに音響特徴ベクトルを予め各単語に関連づけて記憶させておいてもよい。また各単語の音響的な情報が単語の読みのみである場合は、音素又は音節に基づく音響パターンを予め音響モデルとして記憶しておき、さらに各単語の読みを音素又は音節に分解してそれぞれの音響単位の音響モデルの音響パターンを合成する方法によっても、代表単語の音響特徴ベクトルを求めることができる。

次に、インデックス変数  $i$  の値を 2 に初期化する。この初期値は代表単語以外の単語と代表単語との照合を行うために設定するものであり、ステップ S 8 0 2 において元集合の単語リストの先頭要素でない要素を代表単語として選定した場合は、この代表単語と重複しないように他の単語を選択できる初期値を選択する必要がある。

次に、元集合の単語リストの  $i$  番目の要素の音響特徴ベクトルを求める（ステップ S 8 0 5）。音響特徴ベクトルを求める方法についてはステップ S 8 0 3 で説明したので、ここでは省略する。

次に、代表単語と元集合の単語リストの  $i$  番目の要素とを照合するためにそれぞれの音響特徴ベクトルの間の尤度を計算する（ステップ S 8 0 6）。続いてこの尤度が所定の閾値以上となるか否かを検定する（ステップ S 8 0 7）。この尤度が所定の閾値以上となる場合（ステップ S 8 0 7 : YES）は、この  $i$  番目の要素を代表単語のグループに追加する（ステップ S 8 0 8）。その一方で、 $i$  番目の要素に除去マークを付ける（ステップ S 8 0 9）。ステップ S 8 0

9において、 $i$  番目の要素に除去マークを付けるのは、後の処理において、現在の元集合からこの単語を消去するためであるが、ステップ S 8 0 9 で除去マークを付けずに直接元集合からこの単語を消去してもよい。ただしこの場合は、インデックス変数  $i$  と単語リスト中の単語の関係及び元集合に含まれる単語数の調整を行う必要がある。

次に、インデックス変数  $i$  に 1 を加え（ステップ S 8 1 0）、この  $i$  が元集合の単語の個数以下か否かを検定する（ステップ S 8 1 1）。 $i$  が元集合の単語の個数以下となる場合（ステップ S 8 1 1 : YES）は、この代表単語に音響的に類似する単語の探索が単語リストの終端までまだ到達していないことを意味するから、ステップ S 8 0 5 からの処理を繰り返す。しかし、 $i$  が元集合の単語の個数を超える場合（ステップ S 8 1 1 : NO）は、この代表単語に音響的に類似する単語の探索が単語リストの終端に到達したことを意味するから、この元集合にはこの代表単語に音響的に類似する単語はこれ以上存在しないことになる。この場合は、それまでにステップ S 8 0 9 で除去マークを付けた単語を元集合から除去する（ステップ S 8 1 2）。これにより、次の代表単語とグループの選定処理における元集合は、これまで選定された代表単語とそのグループに属する単語を含まないこととなり、単語が複数のグループに重複して属することを回避できる。続いて、ステップ S 8 0 1 の処理に戻り、他の代表単語の選定とその代表単語に音響的に類似する単語のグルーピングを続行する。

一方、ステップ S 8 0 7 において、尤度が所定閾値未満

となる場合（ステップ S 8 0 7 : N O）は、この単語はこの代表単語に関連するグループに属さないことを意味するから、この単語をこの代表単語に関連するグループに追加する処理は行わず、ステップ S 8 1 0 に移って、元集合が

5 含む次の単語の処理の準備を行う。

以上が、本発明に係る音声認識インデックス作成手段の実施例である。

なお、ステップ S 8 0 7 における所定の閾値を大きくすると、ある代表単語に音響的に類似するとしてこの代表単語のグループに属することとなる単語の数が減少すること

10 となり、逆に数多くの代表単語とグループを選定する必要が出てくる。したがって、本発明に係る音響認識インデックス装置における代表単語と入力音響特徴ベクトルとの照合回数が増加することになる。しかしその一方で、この所

15 定の閾値を大きくすることで、グループに属する単語の数が減少するので、各グループに含まれる単語に関連する情報（読み情報やその他の情報）を記憶するために必要となるメモリの量は削減できることになる。

これに対して、ステップ S 8 0 7 における所定の閾値を

20 小さくすると、ある代表単語と音響的に類似する単語の個数が増加するので、一つのグループに属する単語の個数も増えることになり、その結果、代表単語の総数は減少することになる。そのため、本発明に係る音響認識インデックス装置における代表単語と入力音響特徴ベクトルとの照合

25 回数は減少することとなるが、代表単語と入力音響ベクトルとの照合が成功した後のそのグループ内の単語との照合

回数は反対に増加することになる。

さらに、ステップ S 8 0 7 における所定の閾値は、本発明に係る音響認識インデックス装置における入力音響特徴ベクトルと代表単語との照合を行う際に用いる閾値と同じ値を用いた方がよい。そうでないと、音響認識インデックスの作成時には代表単語と類似すると判断された単語が、音響認識インデックスによる検索時には代表単語とは非類似となる、あるいはその逆となることも考えられるからである。

10 以上で、本発明に係る音声認識インデックス作成手段の実施例に関する説明を終える。

以上述べた実施例 1 ～実施例 5 について、各手段はハードウェア・ソフトウェアのいずれでも構成できることはいうまでもない。また、ソフトウェアによって構成する場合には、そのソフトウェアプログラムを記録した記録媒体が必要となる。

#### 産業上の利用の可能性

本発明は上記のように構成されているので、ユーザが音声によって位置情報を検索する場合、ユーザが入力した音声と上記の音声認識登録装置によって予め単語・位置情報関連辞書に登録された単語情報とを照合することで、その単語に関連づけられた位置情報を求めるため、標準パターンを予め登録しておかなくても位置情報を検索することが可能となる。

また本発明は上記のように構成されているので、音響単

位混同行列に基づいて展開された単語も上記単語・位置情報関連辞書に登録されるため、音声入力手段の認識精度やユーザの発声状況によって本来発生した単語とは異なった単語と認識されうる場合であっても、上記単語・位置情報

5 関連辞書に展開された単語と照合を行うことにより正しい位置情報を検索することができる。

また本発明は上記のように構成されているので、大量の単語との音声認識照合を行う場合に、音響上類似する単語との照合のみを行うため、無駄な照合処理を廃することができる。

10

また本発明は上記のように構成されているので、大量の単語との音声認識照合を行う場合に、音響上類似する単語との照合のみを行うため、無駄な照合処理を廃することができる。

15 また本発明は上記のように構成されているので、複数の単語を音響的に類似する単語群に分類することができ、また代表単語も自動的に決定することが可能となる。

また本発明は上記のように構成されているので、任意の複数単語とこれに関連づけられた情報を音声により高速かつメモリの使用効率よく検索することを可能となる。

20

また本発明は上記のように構成されているので、ユーザが登録した任意の単語に関連づけられた位置情報を、認識精度の劣化をカバーしつつ検索可能とする一方で、外部記憶装置が記憶する単語と位置情報の検索を高速かつ低メモリ使用量にて可能とするものである。

25

## 請 求 の 範 囲

1. 音声入力手段により入力されたユーザの発声する音声を音響分析して入力音響特徴ベクトルを出力する音響分析手段と、音節又は音素を音響単位とし所定の音響単位毎の音響モデルを予め記憶する音響モデル記憶手段と、この音響モデルと上記入力音響特徴ベクトルとを照合し関連づけられた音響単位認識候補列を出力する音響単位認識手段と、この音響単位認識候補列とカーナビゲーションシステム本体から取得した位置情報とを関連づけて単語・位置情報関連辞書に登録する単語・位置情報登録手段と、上記音響分析手段により出力された入力音響特徴ベクトルと上記単語・位置情報関連辞書の各音響単位認識候補列の音響特徴ベクトルとを照合して尤度を計算しこの尤度が所定の閾値以上となる音響単位認識候補列に関連づけられた位置情報をカーナビゲーション本体に出力する位置情報検索・出力手段とを備えることを特徴とするカーナビゲーションシステム用音声認識装置。

2. 人間が発声する実音響単位が上記音響分析手段の認識精度のために異なる認識結果として認識される確率を各認識結果音響単位毎に予め記憶する音響単位混同行列記憶手段と、上記音響単位認識手段が出力する音響単位認識候補列の各音響単位とこの音響単位について上記音響単位混同行列記憶手段が記憶する確率が所定の閾値以上となる上記認識結果音響単位とを置換し上記音響単位認識候補列の類似候補を出力する単語展開手段とを備え、



上記単語・位置情報登録手段は、この類似候補をカーナビゲーションシステム本体から取得した位置情報に関連づけて上記単語・位置情報関連辞書に登録する構成とされたことを特徴とする請求の範囲 1 記載のカーナビゲーション用

5 音声認識登録装置。

3. 複数の単語をその読みが類似するものにそれぞれグループ化して予め作成した各単語グループから選択した一つの代表単語とそのグループとの関連を記憶する類似単語インデックス手段と、上記類似単語インデックス手段の記憶するグループ毎の代表単語の音響特徴ベクトルと所定の音響特徴ベクトルとを照合してそれぞれの尤度を計算しそのうち所定の閾値以上となる尤度の代表単語が属するグループの各単語の音響特徴ベクトルと上記所定の音響特徴ベクトルとをそれぞれ照合して尤度を計算し最も尤度が高い
- 10 単語を出力するグループ内類似単語検索手段とを備えることを特徴とする音声認識インデックス検索装置。

4. 複数の単語からなる元集合から一つの単語を代表単語として選択する代表単語選択手段と、この代表単語の音響特徴ベクトルとの尤度が所定の閾値となる音響特徴ベクトルを有する単語を上記元集合から抽出して代表単語と同じグループに含める音響類似単語グルーピング手段と、このグループに属する単語を元集合から除いた単語の集合を上記代表単語選択ステップの処理する元集合としてこの代表単語選択ステップに引き渡す元集合置換手段とを備える
- 20 ことを特徴とする音声認識インデックス作成手段。

5. 上記位置情報検索・出力手段は、請求の範囲 3 又は

請求の範囲 5 記載の音声認識インデックス検索装置を備え、  
上記単語・位置情報関連辞書と外部記憶装置とが記憶する  
単語とその読みと位置情報とをこの音声認識インデックス  
検索装置により検索し出力する構成とされたことを特徴と  
5 する請求の範囲 1 に記載するカーナビゲーションシステム  
用音声認識装置。

6. 上記単語展開手段は、上記類似候補の音響単位ごと  
に上記音響単位混同行列記憶手段が記憶する確率を抽出し  
て類似候補の確率リストを出力し、上記単語・位置情報登  
10 録手段は、この確率リストを上記類似候補とともに上記位  
置情報と関連づけて上記単語・位置情報関連辞書に登録し、  
上記位置情報検索・出力手段は、この単語・位置情報関連  
辞書が記憶する類似単語候補とこの類似単語の確率リスト  
を読み込みこの確率リストの確率が所定の閾値以上の場合  
15 は上記入力音響特徴ベクトルと上記音響特徴ベクトル生成  
手段が出力した音響特徴ベクトルとを照合して尤度を計算  
しこの尤度が所定の閾値以上となる音響単位認識候補列を  
出力しこの確率が上記所定の閾値未満の場合は外部記憶装  
置が記憶する単語とその読みと位置情報を請求の範囲 3 記  
20 載の音声認識インデックス検索装置により検索する構成と  
されたことを特徴とする請求の範囲 5 記載のカーナビゲー  
ションシステム用音声認識装置。

7. 現在位置検出手段と、地図データ記憶手段と、画像  
表示手段と、画像内指示手段と、目的地入力手段と、音声  
25 認識装置とを備えたカーナビゲーションシステムにおいて、  
この音声認識装置は請求の範囲 1 又は請求の範囲 2 又は請

求の範囲 5 又は請求の範囲 6 記載のいずれか一の音声認識装置とすることを特徴とするカーナビゲーションシステム。

1/13

図 1

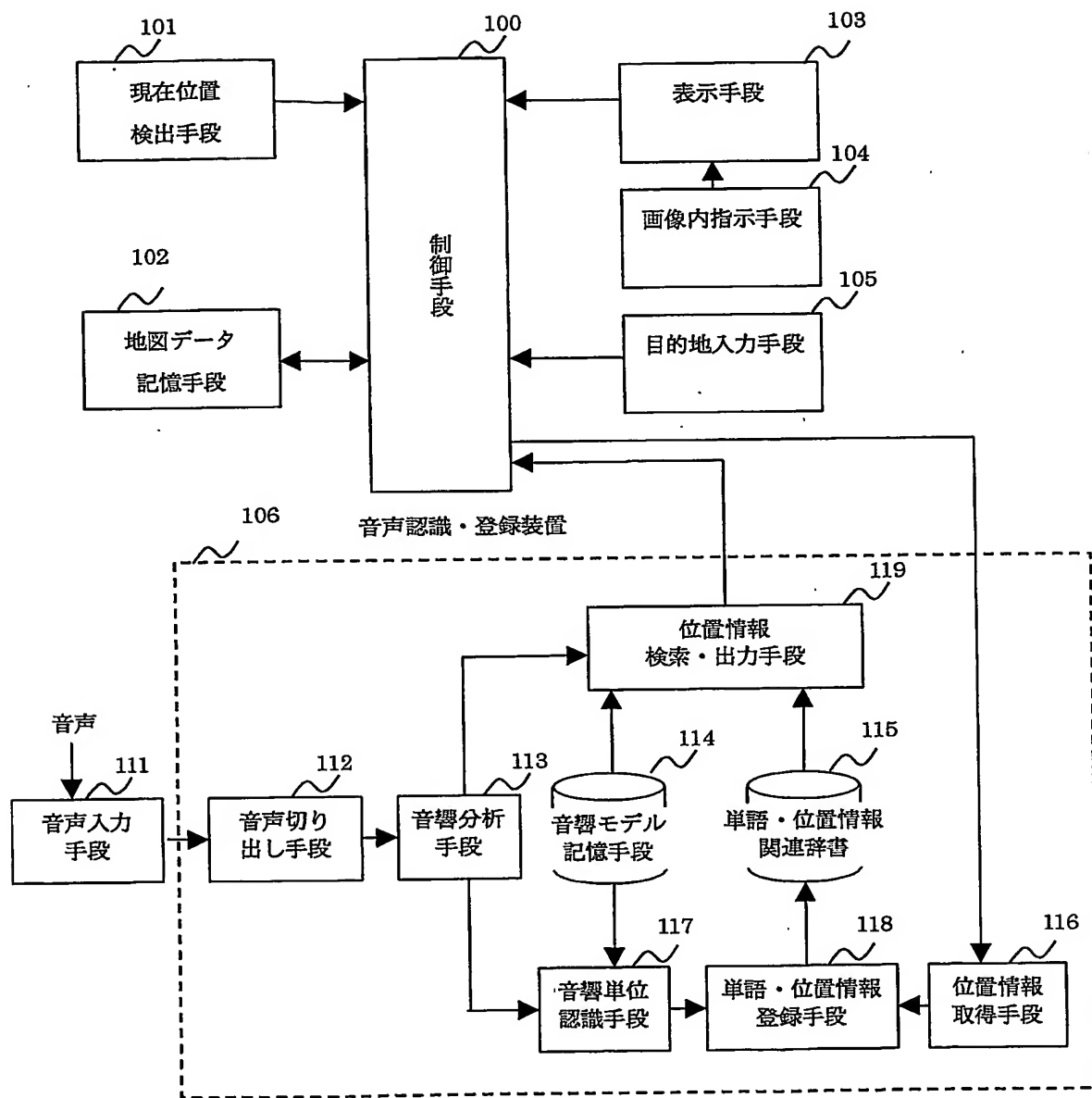


図 2

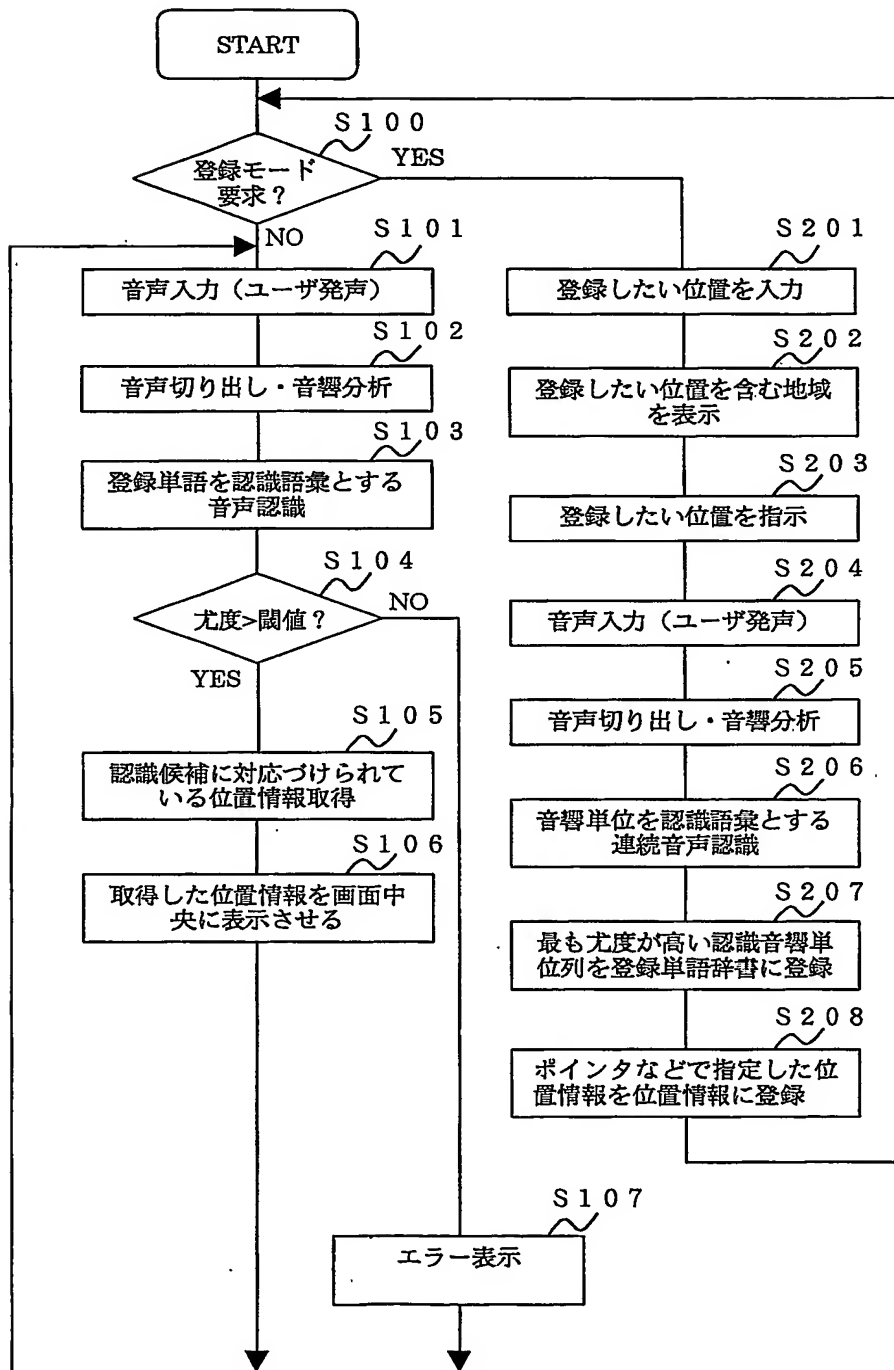


図 3

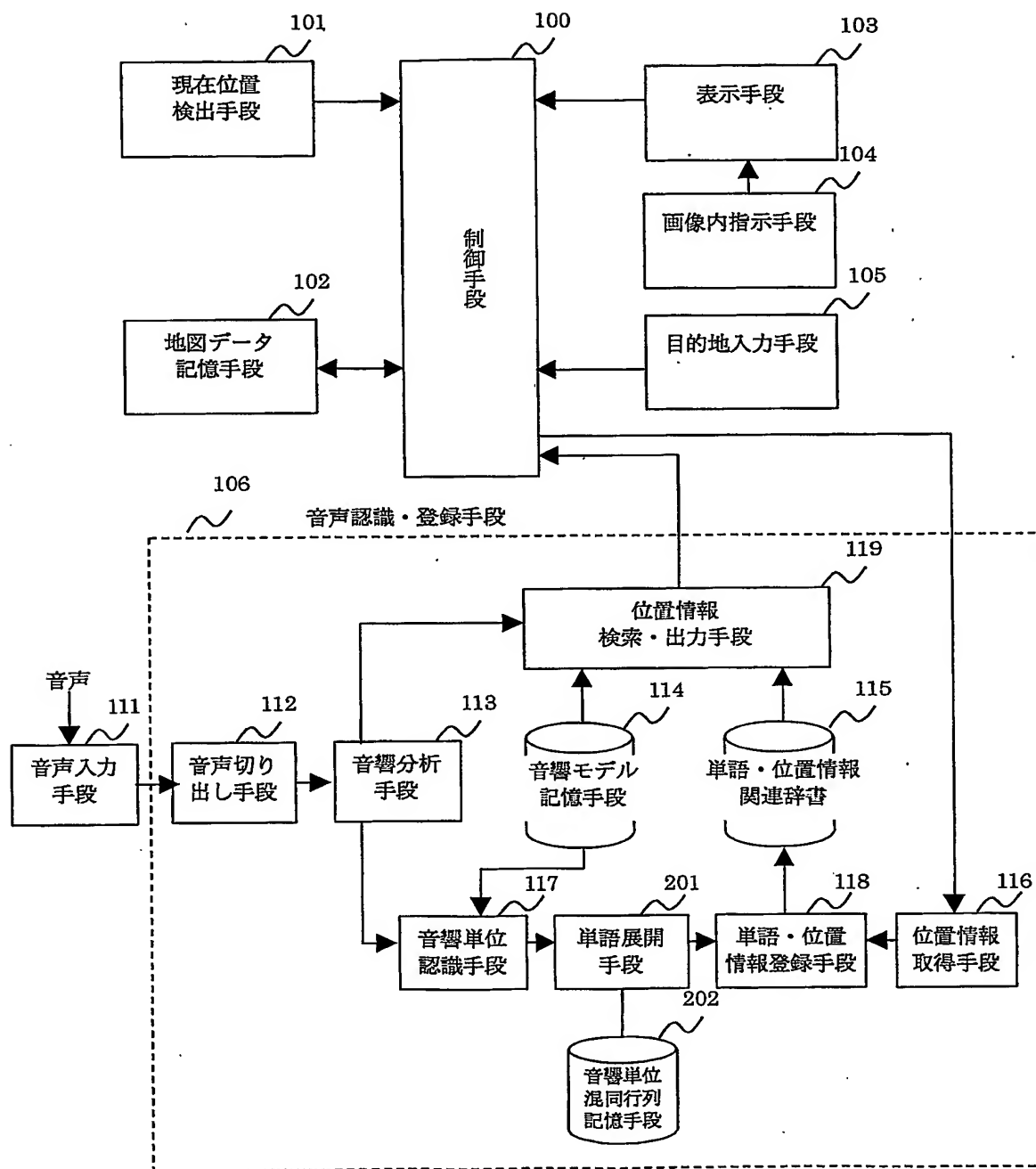


図 4

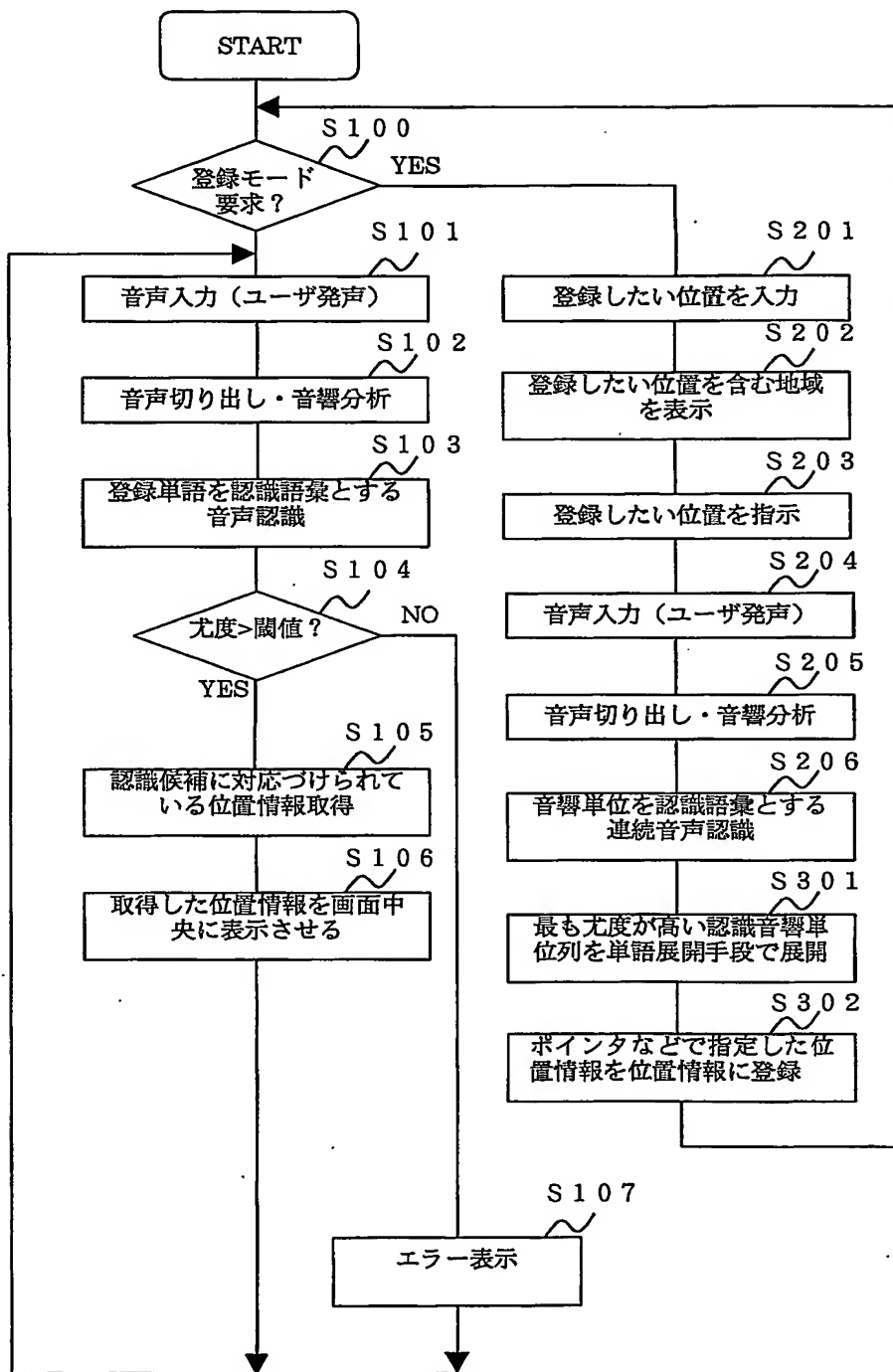


図 5

	...	ka	ki	ku	ke	ko	...	ta	ci	tu	te	to	...	認識結果音節
...	...													
ka		0.92						0.04						
ki			0.95						0.02					
ku				0.91						0.01				
ke					0.89						0.03			
ko						0.99						0.04		
...							...							
ta		0.01						0.93						
ci			0.01						0.98					
tu				0.02						0.91				
te					0.01						0.93			
to						0.03						0.91		
...													...	

実際に発声  
された音節

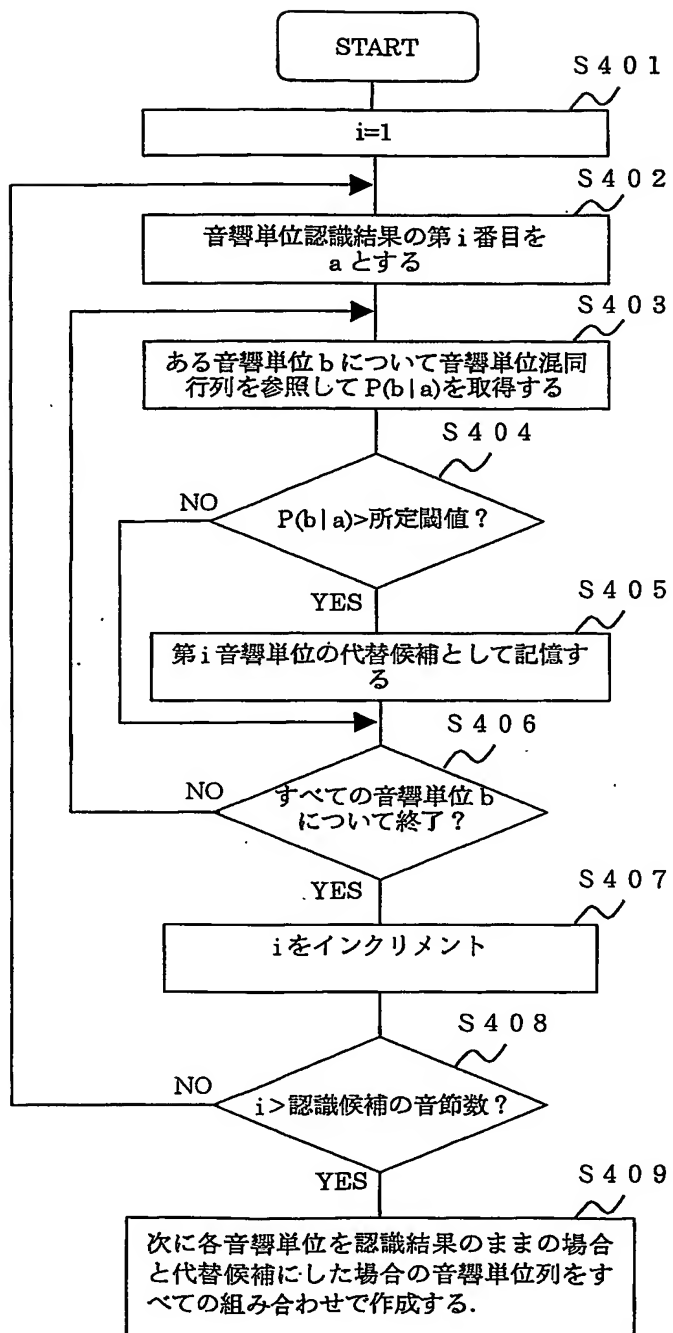
$P(\text{ta}|\text{ka})=0.01$   
認識結果が/ka/でも実際にはtaが発声された確率が0.01

$P(a|b)$  : 音節bが認識された場合に、発声した音節がaである条件付き確率。  
 $P(a|b)=N(a)/M(b)$

$N(a)$  : 音節bとして認識した回数のうち、音節aを発声した回数  
 $M(b)$  : 音節bとして認識した回数



図 6



7/13

図 7

音響単位認識結果  
代替候補

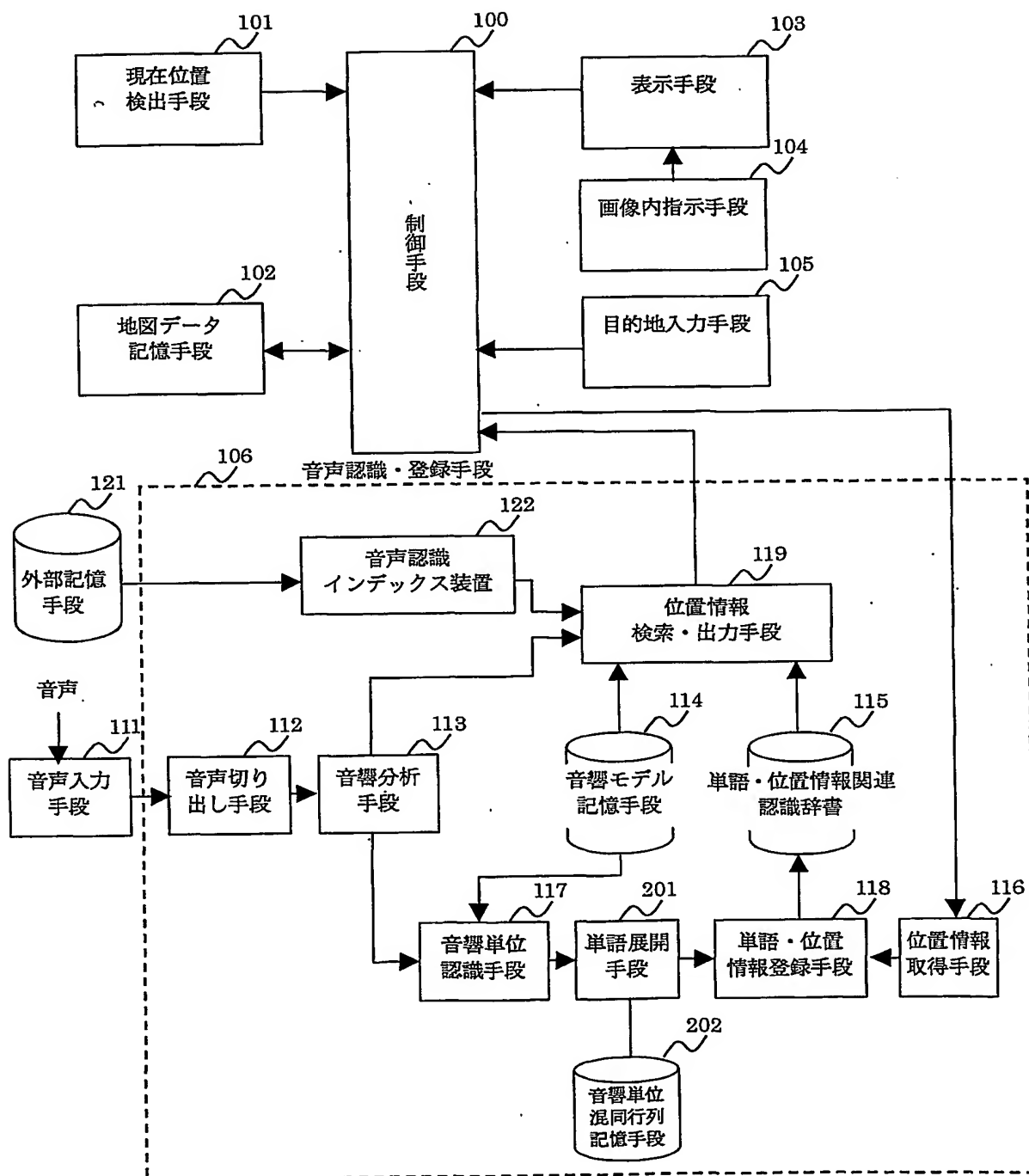
1	2	3	4	5
go	ru	cu	zo	o
		hu	jo	
		u		

音響単位（音節）の代替候補の例

展開した単語

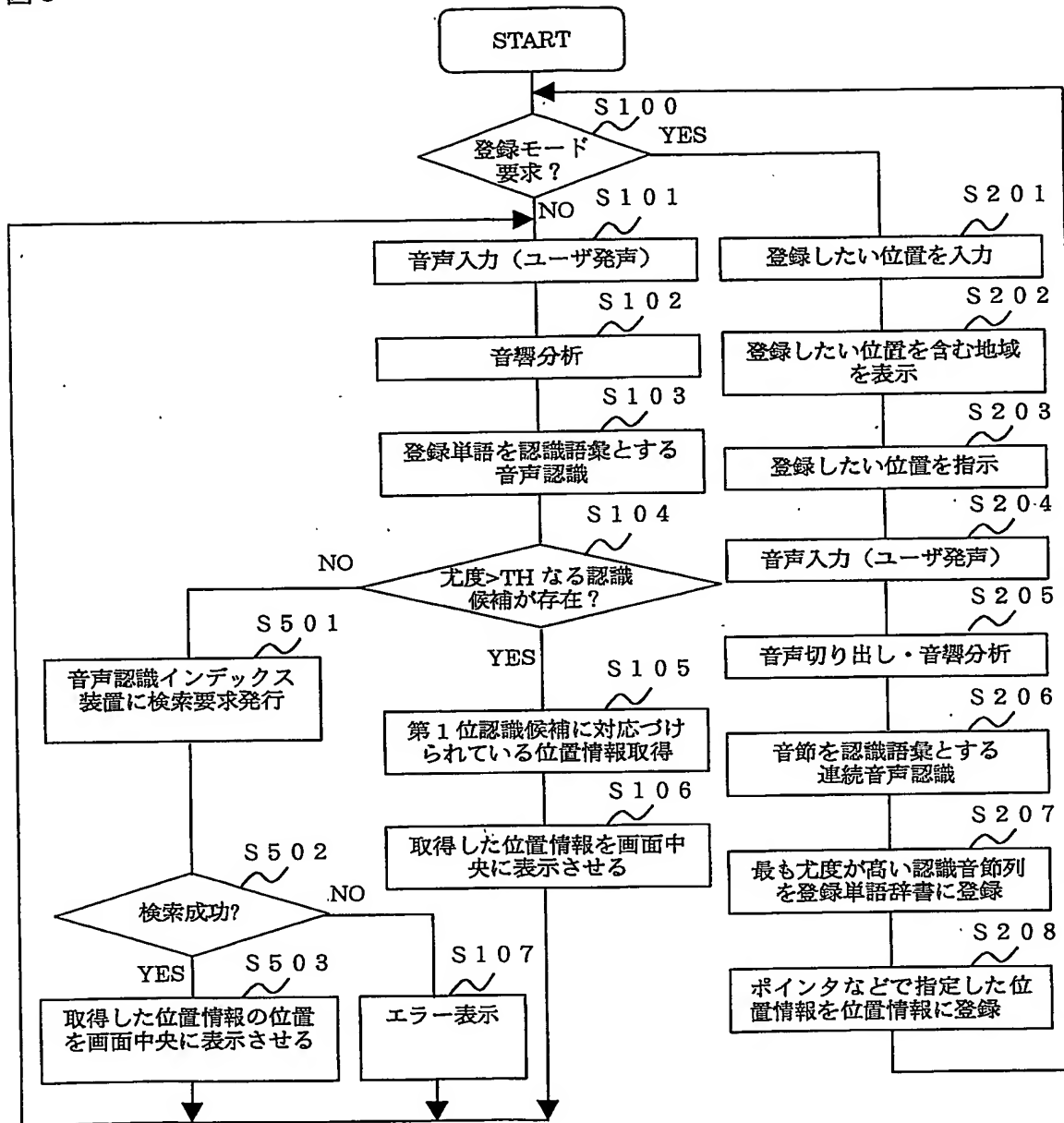
「gorucuzoo」
「goruhuzoo」
「goruuzoo」
「gorucujoo」
「goruhujoo」
「goruujoo」

図 8



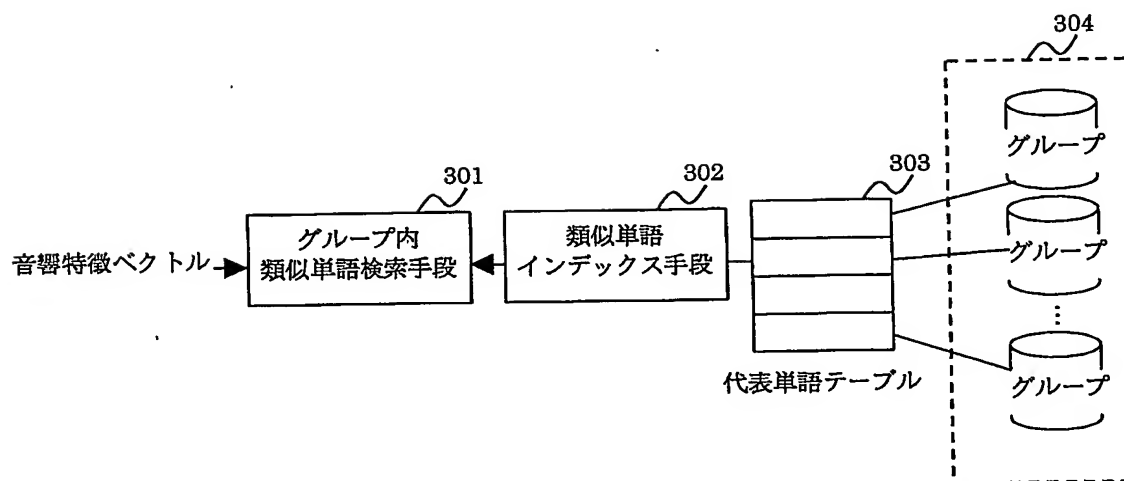
9/13

図 9



10/13

図 10



11/13

図 11

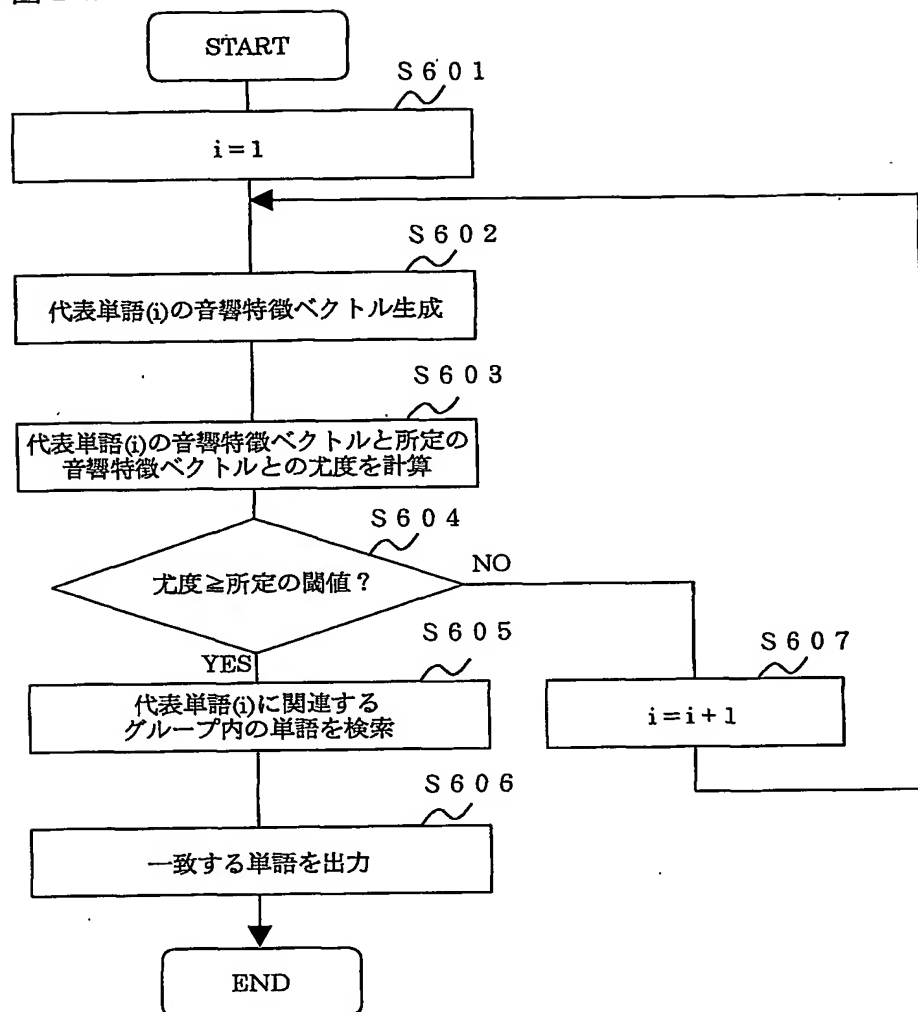


図 1 2

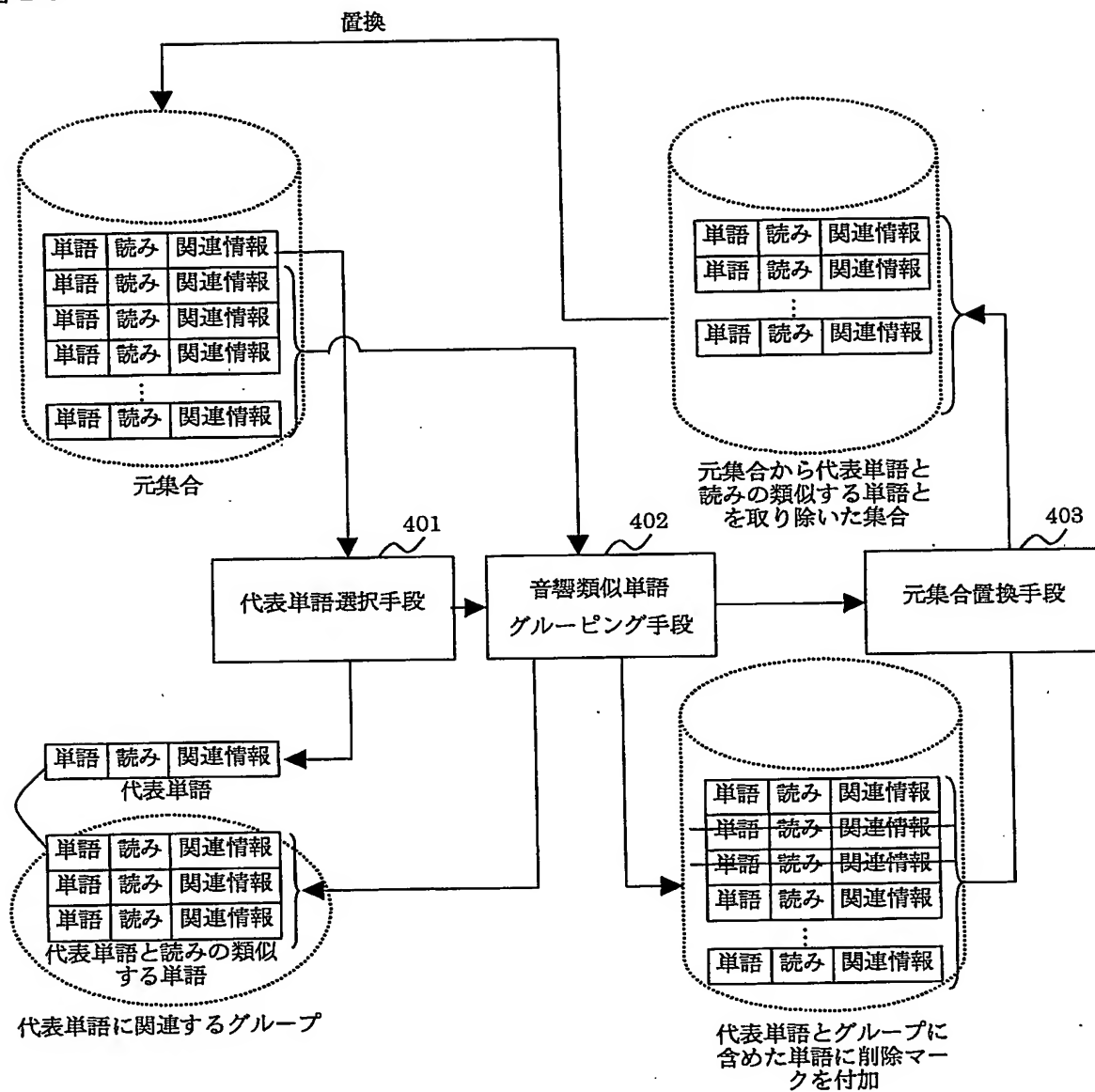
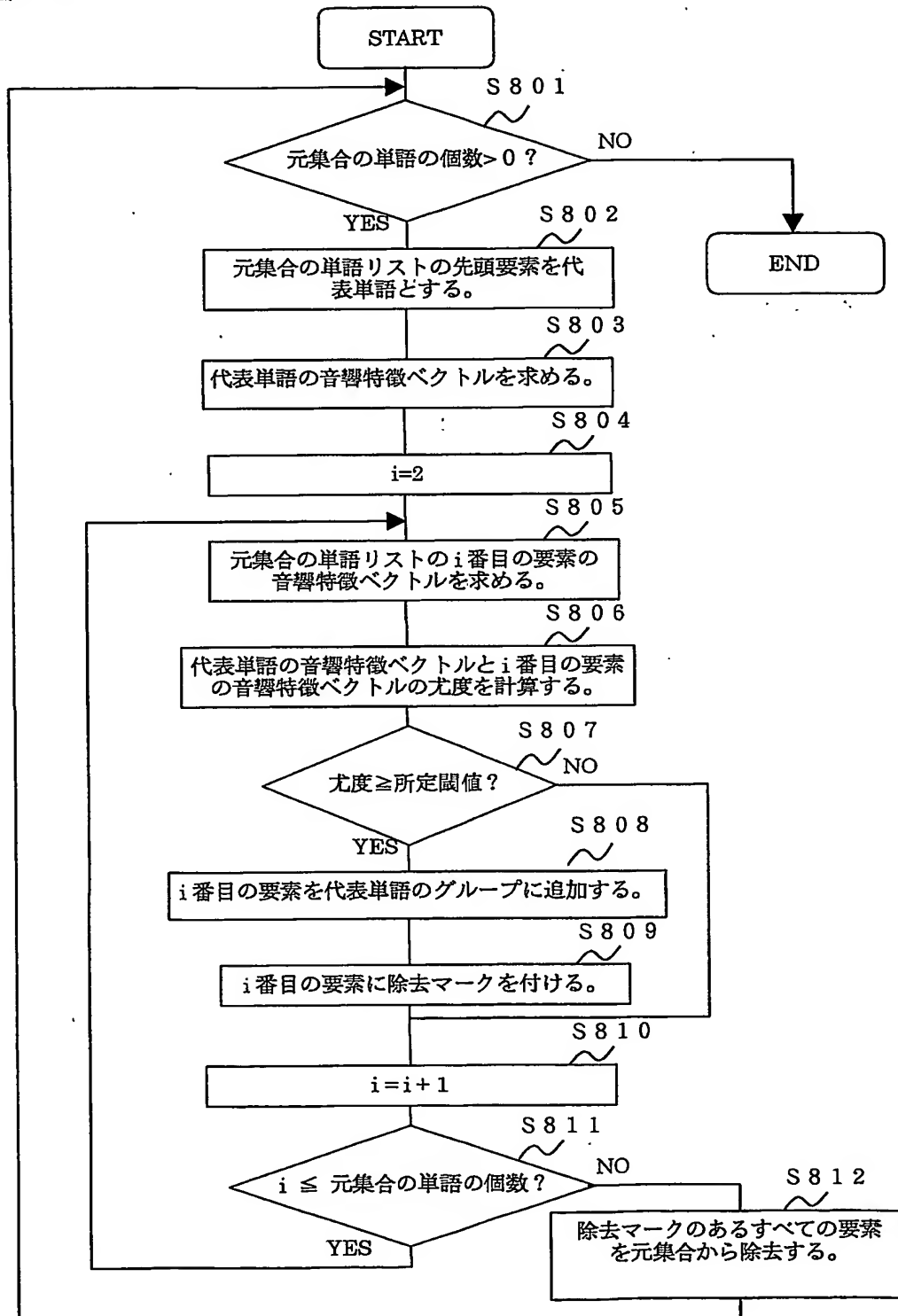


図 13





## INTERNATIONAL SEARCH REPORT

International Application No.

PCT/JP02/03680

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl<sup>7</sup> G10L15/00, 15/06, 15/28, G01C21/00, G08G1/0969

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> G10L15/00-17/00, G01C21/00, G08G1/0969

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1926-1995 Toroku Jitsuyo Shinan Koho 1994-2002

Kokai Jitsuyo Shinan Koho 1971-2002 Jitsuyo Shinan Toroku Koho 1996-2002

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JICST FILE (JOIS)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 8-110790 A (Nippon Telegraph And Telephone Corp.), 30 April, 1996 (30.04.96), Par. No. [0016]; Fig. 1 (Family: none)	1, 5, 7
Y	JP 9-114485 A (Sony Corp.), 02 May, 1997 (02.05.97), Full text; all drawings (Family: none)	1, 2, 5, 7
Y	JP 3-257485 A (Mitsubishi Electric Corp.), 15 November, 1991 (15.11.91), Page 3, upper right column, line 11 to lower left column, line 18; Figs. 2 to 3 (Family: none)	1, 2, 7

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search  
14 May, 2002 (14.05.02)Date of mailing of the international search report  
28 May, 2002 (28.05.02)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP02/03680

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 64-82185 A (NEC Corp.), 28 March, 1989 (28.03.89), Page 1, lower right column, line 16 to page 2, upper right column, line 12	1, 5, 7
Y	JP 62-32500 A (NEC Corp.), 12 February, 1987 (12.02.87), Page 2, upper left column, lines 6 to 20	1, 5, 7
Y	Akinori ITO, Shozo MAKINO "Kakucho RHA-ho ni yoru Renzoku Onsei Ninshiki no tameno Tango Yobi Sentaku", The Transactions of the Institute of Electronics, Information and Communication Engineers, 25 March, 1995 (25.03.95), Vol.J78-D-II, No.3, pages 400 to 408	2, 7
X Y	JP 2-52278 B2 (Fujitsu Ltd.), 13 November, 1990 (13.11.90), Full text; all drawings (Family: none)	3, 4 5, 7
X Y	JP 7-85200 B2 (NEC Corp.), 13 September, 1995 (13.09.95), Full text; all drawings (Family: none)	3, 4 5, 7
X Y	JP 62-134698 A (Fujitsu Ltd.), 17 June, 1987 (17.06.87), Full text; all drawings (Family: none)	3 5, 7
A	JP 11-30994 A (NEC Corp.), 02 February, 1999 (02.02.99), Full text; all drawings (Family: none)	2
A	JP 61-77899 A (NEC Corp.), 21 April, 1986 (21.04.86), Full text; all drawings (Family: none)	4

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl<sup>1</sup> G 1 0 L 1 5 / 0 0, 1 5 / 0 6, 1 5 / 2 8, G 0 1 C 2 1 / 0 0, G 0 8 G 1 / 0 9 6 9

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl<sup>1</sup> G 1 0 L 1 5 / 0 0 - 1 7 / 0 0, G 0 1 C 2 1 / 0 0, G 0 8 G 1 / 0 9 6 9

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1926~1995年  
 日本国公開実用新案公報 1971~2002年  
 日本国登録実用新案公報 1994~2002年  
 日本国実用新案登録公報 1996~2002年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

J I C S T ファイル (J O I S)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	J P 8-110790 A (日本電信電話株式会社) 1996. 04.30, 段落番号【0016】, 第1図 (ファミリーなし)	1, 5, 7
Y	J P 9-114485 A (ソニー株式会社) 1997. 05. 02, 全文, 全図 (ファミリーなし)	1, 2, 5, 7
Y	J P 3-257485 A (三菱電機株式会社) 1991. 11.15, 第3頁右上欄第11行~左下欄第18行, 第2- 3図 (ファミリーなし)	1, 2, 7

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの  
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
 「O」 口頭による開示、使用、展示等に言及する文献  
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの  
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの  
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの  
 「&」 同一パテントファミリー文献

国際調査を完了した日 14.05.02

国際調査報告の発送日 28.05.02

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)  
 郵便番号100-8915  
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)  
榎本 剛

5 C 9 3 7 9

電話番号 03-3581-1101 内線 3541

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	J P 64-82185 A (日本電気株式会社) 1989. 03. 28, 第1頁右下欄16行~第2頁右上欄第12行	1, 5, 7
Y	J P 62-32500 A (日本電気株式会社) 1987. 02. 12, 第2頁左上欄第6行~第20行	1, 5, 7
Y	伊藤彰則, 牧野正三「拡張RHA方による連続音声認識のための単語予備選択」電子情報通信学会論文誌, 1995. 03. 25, Vol. J78-D-II, No. 3, p. 400-408	2, 7
X Y	J P 2-52278 B2 (富士通株式会社) 1990. 11. 13, 全文, 全図 (ファミリーなし)	3, 4 5, 7
X Y	J P 7-85200 B2 (日本電気株式会社) 1995. 09. 13, 全文, 全図 (ファミリーなし)	3, 4 5, 7
X Y	J P 62-134698 A (富士通株式会社) 1987. 06. 17, 全文, 全図 (ファミリーなし)	3 5, 7
A	J P 11-30994 A (日本電気株式会社) 1999. 02. 02, 全文, 全図 (ファミリーなし)	2
A	J P 61-77899 A (日本電気株式会社) 1986. 04. 21, 全文, 全図 (ファミリーなし)	4

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**